

Audio-visual Collections and the User Needs of Scholars in the Humanities; a Case for Co-Development

Franciska de Jong^{◇§}, Roeland Ordelman^{§♣}, Stef Scagliola[◇]

[◇]Erasmus Universiteit Rotterdam
Erasmus Studio & ESHCC
P.O. Box 1738,
3000 DR Rotterdam
The Netherlands
scagliola@zonnet.nl

[§]Universiteit Twente
Human Media Interaction
P.O. Box 217,
7500 AE Enschede
The Netherlands
{fdejong, ordelman}@ewi.utwente.nl

[♣]Netherlands Institute
for Sound and Vision
P.O. Box 1060,
1200 BB, Hilversum
The Netherlands
rordelman@beeldengeluid.nl

Abstract

The aim of this paper is to reflect on the factors that impede a clear communication and a more fruitful collaboration between humanities scholars and ICT developers. One of the observations is that ICT-researchers who design tools for humanities researchers, are less inclined to take into account that each stage of the scholarly research process requires ICT-support in a different manner or through different tools. Likewise scholars in the humanities often have prejudices concerning ICT-tools, based on lack of knowledge and fears of technology-driven agendas. If the potential for methodological innovation of the humanities is to be realized, the gap between the mindset of ICT-researchers and that of archivists and scholars in the humanities needs to be bridged. Our assumption is that a better insight into the variety of uses of digital collections and a user-inspired classification of ICT-tools, can help to achieve a greater conceptual clarity among both users and developers. This paper presents such an overview in the form of a typology for the audio-visual realm: examples of what role digital audio-visual archives can play at various research stages, and an inventory of the challenges for the parties involved.

1 Introduction

The aim of this paper is to reflect on the factors that impede a clear communication and a more fruitful collaboration between humanities scholars and ICT developers. One of the observations is that ICT-researchers, who design tools for humanities researchers, are less inclined to take into account that each stage of the scholarly research process requires ICT-support in a different manner or through different tools. Likewise scholars in the humanities often have prejudices concerning ICT-tools, based on lack of knowledge and fears of technology-driven agendas. If the potential for methodological innovation of the humanities is to be realized, the gap between the mind-set of ICT-researchers and that of archivists and scholars in the humanities needs to be bridged.

At present the academic community within the humanities can be roughly subdivided in three categories that each relate to the emerging paradigm known as ‘e-research’ and the role of ICT in their field of research in a different manner. The variation among the categories is gradual rather than absolute. On the one hand of the spectrum are the scholars who are deeply engaged and actively support the establishment of e-research. On the other end of the spectrum is the gradually decreasing group of dissenters who are reticent about the ‘hype’ and who are strongly dependent on young assistants to take those steps in the digital world that are inevitable. In the middle are those who are interested in taking up new methods, but just lack the training and resources to engage in e-research, and those who pick and choose on an *ad hoc* basis what they can handle and consider useful.

The first group of scholars is typically involved with initiatives aiming at the furthering of the technological

basis of e-humanities and the dissemination of results of successful multidisciplinary collaboration. Among this type of initiatives are conferences tagged as ‘Digital Humanities’, e-humanities, or computational humanities, and projects taking place in the context of EU-wide programmes such as CLARIN¹ or Europeana². This engaged and supportive category of humanities scholars understands and confirms the importance of multidisciplinary ICT/humanities collaboration and develops ideal-typical agendas for the future. Yet what is often underestimated in the agenda setting is the need to reach out to scholars who are less engaged with e-research. The idea that time will solve the problem because the ‘digital born’ generation will eventually hold the leading academic positions underestimates the impact of the lack of specific skills among younger generations of students and researchers. Knowing your way on Facebook and Google does not imply sensitivity to the potential of dedicated search engines for scanning important primary sources. Nor is the role of digital data curation and annotation for archiving and reuse of sources integrated in present day academic curricula.

The argument here is that in order to connect to this group it is important to conceptualizing e-research as an intervention in current practices. In the next section we will present arguments to underline this necessity.

2 Tension between the ‘tribes’

2.1 The role of ambitions and temptations

As Beaulieu&Wouters (2009) have stated, e-research needs to be understood in sociological as well as

¹ <http://www.clarin.eu/>

² <http://www.europeana.eu/>

epistemological terms. The fast development of e-research has a cost in the sense that it ‘disrupts the existing fabric of social relationships that carry knowledge creation’. In addition to the sociological and epistemological point of view we would like to underline the anthropological perspective: scholars are required to change engrained rituals of communication, of analysis, of language, of ranking, and to internalize the language of a new ‘ICT-tribe’, which promises a future practice of increased efficiency, relevance and novelty. Yet, according to a recent Research Information Network Report (RIN, 2011), scholars are only interested in tools that facilitate or simplify an existing practice. The survey carried out as basis for the RIN-report indicates that the assumed indifference of humanities scholars towards digital tools and resources is a deliberate choice: they avoid which is not perceived as being useful. Funds for e-research intended to stimulate cooperation are effective means to overcome hesitations. Yet, this approach runs the risk of leading to grant-driven research proposals phrased in ‘visionary’ terms to convince financiers of a successful outcome. Often only once the funding is granted and the joint project work has officially started, the parties involved gradually become aware of divergent expectations.

It is widely assumed that the best way to circumvent or minimize this kind of complication and frustration in multidisciplinary work is to adopt some model of co-development: a way of collaboration in which all parties actively participate in all stages of the work plan, including the design of the work plan, under the assumption that this model is to the benefit of all parties. In the domain of e-humanities the benefits for the users are supposed to be obvious (see below). But it is not always acknowledged that although ICT researchers and developers find the humanities case very challenging for its inherent complexity and diversity, it is often hard for them to adhere to the principles of co-development. We think it is worthwhile to analyse some of the obstacles for a collaboration that is mutual beneficial more thoroughly in order to implement co-development tracks more effectively. In the next section we will therefore describe the dynamics of multidisciplinary collaboration in more detail.

2.2 The dynamics of collaboration

Humanities researchers can hardly be indifferent to the promise of innovative tools for the support of content exploration and content annotation. Both are key elements in their daily research practice and as such can be considered the alpha and omega of their analytical and comparative work. The perspectives of widened data coverage and control, and of generating more consistent metadata as well as richer and flexible metadata structures, are all very attractive. But the first and main concern of humanities researchers is and will be the research question that drives the exploration for relevant and accessible sources. This is not always clear to ICT-researchers, whose point of departure and challenge is opening up the archive, the collection or the library.

So, whereas the historian’s is driven by a question of which the answer may or may not be enclosed in the archive, the challenge for the ICT-specialist is to find technical solutions to increase the chances of finding exactly the right information within the archive.

As described above, in multidisciplinary collaboration, a clear distinction and mutual understanding of the aims of each partner is crucial, and the chances that the expectations need to be adjusted during the course are high, as are the chances that the certain assumptions need to be corrected. Just to give some examples of likely sources of confusion in settings where dedicated search technology is the common objective: (i) finding more in less time may not be a goal in itself for humanities researchers, while for ICT researchers by default increasing the efficiency of systems is part of the job, (ii) for humanities researchers, a deep engagement with the primary texts that are the context of their research is necessary for the identification of useful threads within the content. Many researchers still consider printing part of the primary texts and spreading sheets on their desk or floor, as an essential step for grasping the essence of the content. And the stage of text composition working from a screen can feel as if one is playing chess with a minimal view on the chessboard. So for a considerable group of researchers the widely appreciated availability of digital content does not always fully replace printed content.

The examples underline that alignment of the expectations for co-development is a *conditio sine qua non*, and should be monitored during all stages of the process. In the next subsections we will go deeper into some likely causes for failing collaboration.

2.3 Moving targets

The typical ICT cycle starts with a user study and the analysis of the requirements that need to be fulfilled in a scenario of use or for a specific task for which users seek technology support. The requirements specification is the basis of a design stage that then leads to the development and testing of a prototype, and the evaluation of the prototype in some context of use. The outcome of the evaluation can be the starting point for an iteration of this process.

In general there can be several reasons why an iteration is needed, but in the humanities case, specific situations may emerge: (i) the users that inform the requirements analysis are unaware of the possibilities and limitations of technology and during the project it is discovered that the requirements collected are not matching the technology baseline and the potential enhancement, and (ii) due to the novel functionality the humanities scholar progressively gets a deeper insight in the collection coupled to the prototype; as a result of the deeper understanding, their research question or the leading hypothesis is re-formulated. This may ask for adjustments of the researchers’ requirements that in turn have to be reflected in changes of the design. This shift in perspective may not easily synchronize with the standard development stages of the ICT researcher.

2.4 Experiments *versus* product

When ICT researchers and humanities scholars join forces in order to adapt the state-of-the art in technology to the requirements of scholarly research, the scholars are often not aware of the fact that they have joined an ICT experiment and that they are not the partners of a commercial software developer who gathers requirements and just does the trick. Beaulieu & Wouters strikingly characterize these coalitions as a ‘parade of prototypes’, applications developed within the framework of a particular project, on an experimental basis, but aimed at the realization of a more generic tool that can serve a broader area of scholarly communities. Often it is not clear whether these ICT experiments will be supported by a next round of funding or by a take-up of the resulting technology by the commercial parties that provide and operate the systems that are used in the archives, libraries, and data networks for the curation of (research) data collections. (See also the discussion in Section 3.4)

Involving commercial parties at an early stage in these coalitions that have invested in carefully monitoring methodologies of various groups of researchers in order to develop adequate selling products, can contribute to a viable long-term investment. (Beaulieu & Wouters, 2009). This may also increase the chances that tools that are already used and appreciated by scholars, either or not from a commercial supplier, are not ignored, and that the feasibility of coupling of existing and new functionalities is taken into account in the analysis of requirements.

2.5 Impact on research?

Among the noteworthy observations in the RIN-report mentioned above on the changes in research practices in the humanities is the following paradox: in the evaluation of novel text search tools, scholars emphasize on the one hand the ease and speed of access to information, and the ability to conduct research that was impossible before keyword search, but on the other hand, they indicate that this has not led to the emergence of novel research agendas and paradigms. Many interviewees describe their work as similar to what they did before when they had just printed documents available. Assuming that the co-development model could lead to a more fruitful outcome, ICT-researchers will need to gain more in-depth knowledge of the diversity of methodology and disciplinary-specific assumptions within the various academic communities. This means acknowledging and understanding the specific descriptive, performative and solitary character of a considerable part of present-day academic research in the humanities, unless one wishes to focus only on younger generations who grew up in a paperless world and who think they know how to get the best out of it.

2.6 Lack of training and unclear terminology

In general, humanities researchers are poorly informed about the iterative design procedures underlying ICT development and the role that they as users could and should take up in the stages of this process. Training should be supplied for tuning their workflow to a version

that is adjusted to the new possibilities. This applies to additional courses for established researchers as well as for more basic courses for students. Another aspect that needs attention is the potential confusion due to differences and domain-specific subtleties in vocabulary. Discussions on how to set up a research proposal or on the benefits and disadvantages of a specific approach are often blurred by the sheer fact that certain terms have a different meaning and association in the ICT and humanities realm. A concept such as ‘metadata’, refers to concise information on a specific object for a humanities scholar, it is related to content, it gives him or her the opportunity to select. For an ICT-researcher metadata has a far more extended meaning, it is not only about content, but also about format, technical aspects, and results of processing data. The same accounts for the term ‘annotation’. It originally means making a note while reading a text, but as the content of digital archives nowadays can be enriched with automatic annotation, crowd annotation and many other forms, an ICT-researcher has a much broader scope in mind when using this term.

Early adopters of e-humanities could play a more active role in clarifying the importance of some basic training in these issues with at least a common vocabulary.

3 Stages in scholarly e-research

In the remainder of this paper we will illustrate the diversity of requirements at different stages of the research process with audio-visual collections as a case study. This is a deliberate choice as in our view the differences in ‘mindset’ between the two groups of researchers manifest themselves strongly in this domain: there is an emerging interest of scholars to explore audio-visual collections, ICT is already explicitly present in large-scale preservation and access initiatives, and last but not least, the challenges with respect to the analysis and representation of audio-visual content for access purposes are attractive for ICT researchers. The characteristics of the audio-visual domain are elaborated on in the first section, where the connection is made between the nature of audio-visual sources, the potential of technology and the limitations of manual handling of data. In the following subsections the various stages of the research process and the relevance of ICT are described, both in text and integrated in Table 1, a schematic overview which is presented at the end of Section 3. To further illustrate the stages, two scenarios have been developed in which a linguist and an historian pass through the various stages, while at each stage it is indicated how in which way they could deploy ICT-tools. The scenarios are outlined in Tables 2 and 3, also at the end of this section.

Although the schemes and scenarios may at first sight seem redundant to ‘converted’ e-scholars, what must be taken into account is that they are meant to lower the threshold for first time users, and sensitize the ‘insiders’ for the more basic requirements of clear communication. In our view this will ease the possibility of finding joint solutions for technical and conceptual obstacles on very specific levels, such as protocols and standardization.

3.1 A/V content & scholarly research

The general assumption is that the large amounts of audio-visual (A/V) content in (audio-visual) archives form a rich basis for various types of scholarly research in the humanities and social sciences such as history, literature, linguistics, political science, sociology, communication studies and cultural studies. To serve as a resource for research an effort is needed to ensure that digitized material is encoded and presented in ways that suit specific methodologies of scholars. After all, A/V content is in itself unstructured, and –being composed of pixels and samples– typically lacks representation formats that can easily be scored or collated, as is the case with OCR-ed text from a book. On the other hand, the multi-semiotic nature of A/V adds dimensions for inquiry that do not exist in written text (Goldman et al., 2005). For example, the interpretation of an opinion or argument in an audio-visual source offers the opportunity of taking the variety of semantic dimensions which are expressed on this multimodal carrier into account, such as the intonation of a speaker's voice, the facial expression, the language that is used, the accompanying video shots, or context in terms of social-cultural discourse (De Jong et al., 2008).

The traditional approach towards bridging the semantic gap between the low level features of audio-visual content (the pixels and colour samples) and the information needs of users has been to let professional archivists provide manual annotations, commonly known as metadata. It is evident however that with the increasing data quantities (due to retrospective digitization and the general increase in A/V production volume) there is a limit to the portion of annotation that can be covered by the manual approach. In addition to aspects of data quantity, the required quality of the metadata puts pressure on available manual resources. Traditionally, archival descriptions focused on the level of the document: typically some technical metadata, a title, a list of people occurring in the video, a short summary, and ideally also some labels extracted from a controlled vocabulary or thesaurus. However, recent user studies as the ones conducted by Huurnink (2010) and Van den Heuvel (2010) suggest that users are specifically interested in *fragments*: archival footage of the pre-war city centre of Rotterdam, a shot of a bee on a nose, a quote of some specific minister, a topic being discussed, or the most exciting parts of a soccer match. Occasionally, archivists may label fragments in the video, for example by noting down the start-times of the topics appearing in a news show, but the manual labelling of every fragment that is likely to be interesting for some user is obviously unfeasible.

One approach towards bridging the gap between fragment-level information needs and archival description capacity is the use of (semi-) automatic annotation strategies focussing on the automatic analysis of either the audio track or the visual channel. Typically such analysis tools produce labels describing the content, such as transcripts of the words spoken, the identity of the speaker, or a visual concept that appears, and –very importantly in the context of fragment level search– time-codes indicating exactly *where* the

labelled item occurs in the video (Ordelman et al., 2009). Optionally, with these labels and time-codes also 'confidence scores' can be provided. These can be generated by the analysis tools and are indicative for the level of certainty that the label is indeed correct.

The past decade a vast amount of research has been invested in technology for automatic generation of annotations, search, analysis and interpretation of AV-content. Examples are speech recognition technology for converting speech in the audio channel to text, visual analysis for detecting occurrences of visual concepts such as people, or objects such as a boat, or a tree in the video stream, and speaker identification based on the automatic analysis of the audio and/or video channel. In addition, opportunities to let 'the crowd' (ranging from communities of knowledgeable experts to interested amateurs) participate in generating annotations have been studied and put prudently into practice. In an archival context, automatically generated annotations and content descriptions coming from external users are all referred to as metadata and just like the good-old metadata generated by archivists employed by the keeper of the content, can be deployed for building indices to search the archive.

It should be noted here that the 'trustworthiness' of automatically generated annotations and annotations obtained using social tagging mechanisms, is becoming an important issue in the archival domain as it touches upon the principals of authority and quality an archive adheres to. Annotation technology produces varying levels of accuracy, depending among others on the complexity of the task and the characteristics of the data. For example, automatic speech recognition yields far better (nearly perfect) annotation results for news broadcasts than for interviews with elderly people talking in dialect (Ordelman et al., 2009). Furthermore, annotation technology may be limited with respect to the instances they can label. For example, the performance of tools for the detection of visual concepts (such as 'a tree', 'an airplane', 'outdoor') can offer adequate results for a relatively small number of concepts only. Automatically generated metadata should be regarded as tags within certain margins of trust, rather than as perfect annotations. The same holds for the quality and precision of tags produced by 'the crowd'. Dealing with different levels of trust is a well-known consequence of using annotations coming from other sources than archivists (see also Ceolin et al., 2010) and has implication for the use of such novel annotation technologies.

3.2 Preliminary archival search

The effort put in disclosing the archival domain – one of the primary places where researchers look for suitable sources – is an investment on a general level. Typically the perspective is not geared towards specific disciplinary audiences and their needs. The first steps in content exploration by a researcher often come down to searching for material. Research starts with search and this can be just browsing as a general interest, purpose-driven, (e.g., checking details, searching for

complementary sources), item-oriented (e.g., finding the first edition of a certain text), or directed towards a collection as a whole, in which case an entire dataset is the focus of attention.

The archive or library is the context for this first stage and support is therefore expected for the identification and retrieval of (links to) items or documents that very precisely match an information need.

Note that the archives themselves can be the result of research projects. For example, interview collections can be created in the context of oral history. Before the digital turn, tapes and diskettes with the recordings were typically neglected or forgotten after the manual elaboration that resulted in summaries and transcripts. Nowadays it is possible to process, index and store them. In order to enhance the chances that others identify such recordings as a useful source for research, annotation and storage protocols should ensure that they can be found.

Searching collections can be performed at home or at the office, according to the state of the digitization of the archive or library. Yet the absence of a human intermediary, a specialist that can be consulted face to face, has consequences for the organization of the searching of the content, for the interface, and for the necessary background knowledge of the consultants. Column A in Table 1 below gives an overview of the various tasks and roles of the different professions involved in the archival search stage.

3.3 Content analysis of a specific data set

After the relevant materials have been identified, in the next stage of a research project the focus is mostly on the exploration of the collected materials, the ordering, comparison and analysis, and the documentation of the interpretation. In the audio-visual case ICT can offer support in this stage by providing functionality that allows a researcher to get deeper involved with the materials through watching and listening. This exploration stage may generate new ideas and perspective. At the same time the exploration may lead to new searches and inquiries.

To be able to decide on the suitability of a data set for further research, a user is partly dependent on the available means to do the scientific digging. Traditionally, the analysis is done manually by browsing through the data set and play, listen and watch the audio-visual content. When annotations exist –either manually or automatically generated- the process of digging up interesting fragments can be accelerated. However, next to annotations, also tools are needed that can take annotations as input and provide means for visualization, compression and aggregation of the data to allow the researcher to make sense out of the (often still vast amounts of) data. Ideally, the researcher is able to select from a variety of highly specialized annotation tools that provide him with annotation options geared towards his research hypothesis.

When creating a tool that aims at helping to answer a specific research question by processing the material in a specific way, technology specialists and researchers in

the humanities have to engage in a joint process, the co-development of the tool, taking into account the specific set of practices and conventions coonly adhered to within the discipline. In order to increase the chances for a serious return on investment it is relevant to have a good estimation of what the added value could be in other domains. This requires an insight into the diverse methodologies. For instance, if a psycho-linguist and an ICT-researcher create a tool that links facial expression to words and sentiments, the question to be raised is whether this can also be applied to similar tasks conducted in other subfields that come with slightly varying requirements. Column B in Table 1 captures some crucial ingredients of content analysis stage.

3.4 Presentation and dissemination

A third stage where ICT is applied is that of the presentation of research results, after the analysis has been completed. In the digital realm it has become feasible to link annotations that capture the results of an analytical step to the data on which they are based. Instead of a printed book, one can produce a digital publication with links to video- and audio-content. This meets one of the basic principles of scientific research; the possibility to check outcomes of research by getting access to the primary sources. But besides sheer scientific added value, this application of ICT enlarges the educational potential of audio-visual sources by making the final outcome of a research livelier and more appealing. One may even open up the possibility to engage the audience (scientific peers or general public), by offering them the option to provide feedback through suitable modes of interaction. Of course the option of feedback and annotation entered by the crowd through some web applications can only be of added value if an investment is made in monitoring the result in order to guarantee the quality of the archive and its sources. Column C in Table 1 captures some crucial ingredients for result presentation/dissemination stage.

3.5 Curation

The actual final stage of the process, imposed by the very fact that all sources are now in digital form, is the long-term preservation of the data, and of the results of the investigation that has been carried out. Especially audio-visual materials that have been processed with digital tools are not the kind of research result that can be stored in a cupboard; they should be deposited in a trusted digital repository. Ideally the depositing of material should be in line with emerging standards for Open Data, as this would allow the data and annotations to be reused by scholars with similar interests. As services of archives, libraries and universities are paid with public money, it is reasonable to expect researchers to adhere to principles of sharing, provided authorship issues and publishing rights can be covered by adequate models for the acknowledgement of the academic effort that goes into the creation of research data.

Table 1	A. within the archival domain	B. within a specific data set	C. result presentation	D. data curation
What researchers (should) do/need	<ul style="list-style-type: none"> • exploration • browsing • search for complementary information • search for suitable data set • search for original/un-edited versions 	<ul style="list-style-type: none"> • discern of patterns • visualisation of patterns • comparison of patterns • peer-to-peer collaboration 	<ul style="list-style-type: none"> • digital platform to present results • tools/expertise for creation of enhanced publications (with links to data) • training in attributing metadata and relevant standards 	<ul style="list-style-type: none"> • understanding of the need for long term • training in digital referencing reference to the project in repository index
Available ICT-applications	<ul style="list-style-type: none"> • digital archives • search at document-level • search at fragment-level based on time-labelled A/V • (semi-) automatic annotation and crowd-sourcing • linking within and across datasets 	<ul style="list-style-type: none"> • tools for automatic content analysis in speech, text and (moving) images • workspaces for digital collaborations • platform for editing • visualization tools 	<ul style="list-style-type: none"> • enhanced publications • liquid publications • visualisation tools • persistent identifiers • content play-out (streaming) • digital rights management 	<ul style="list-style-type: none"> • repository to deposit digital data and processing tool(s) • long-term data preservation and migration • curation of specialised (annotation and/or presentation) tools
Partners in co-development, additional to scholars and ICT developers	archivists, librarians, data managers, legal experts, commercial software developers	methodologists, commercial software developers	publishing houses, archivists, editors of journals, editors of websites	archivists

Table 2: Scenario I

A. A historian interested in the way Germany and Germans have been represented on Dutch television in the post-war period from 1945 to 2000 and browses through the audio-visual archive to look for suitable sources. On the basis of meta-data and some previous research on this topic he decides not to focus on political and current affaire programs, but on German detectives, when and how did these series become popular and how does this relate to German-Dutch relations, what did the Dutch think of hearing the German language on television? Gradually it becomes clear that the archive provides only half of the sources for this research. With the help of data linking the historian is able to find information in old broadcast magazines and newspapers with comments from readers on these series. He also has access tot he digitized archive of the broadcasting company, where he has found the correspondence on the negotiations with regard to buying this series.

B. He decides that he wants to compare the references to Germany and Germans in Dutch in printed and audio-visual form, both from before the series were introduced and from after some years. As the audio-visual archive offers the possibility for text search and speech search, this approach is feasible.

C. The outcome of this research is a digital publication with references to the data that have been processed. This is presented on the website of the university and in a journal for media history. The company that broadcast the series is interested in reintroducing the series and asks viewers to send their reminiscences with regard to feelings about Germany and Germans at the time. This feedback is assessed for suitability and the best stories are linked to the publication.

D. The publication and the data are deposited in a digital repository, with references to the tools which were used.

Table 3: Scenario II

A. A linguist is interested in tracing the variation in intonation of anchormen of current affairs and news programs over time; she or he browses through the catalogue and content of the archive and finds 4 or 5 suitable collections. On the basis of the metadata she can conclude that program A and B are suitable, as the archive covers 20 years of this program which are all digitized. In addition the archive offers the possibility to search the material with speech retrieval. She manages to collect the necessary range of 'speech acts' to perform a specific search.

B. She can process her collection with a tool that can analyze the differences in intonation and has been developed and provided by a linguistic institute. They can provide support and advice.

C. The outcome of the research is presented on the website of her university with a link to her data, the same is done on the website of a European project which is the context for this research, possibly with a translation tool. She also presents her project to a peer reviewed online journal, and the experts can check her analysis by getting access to her data, they add comments to her publication and after a few weeks she gets feedback which can be included in the final version which is automatically updated in all digital spaces.

D. She makes sure that the data and the tool which she has used are deposited in a digital repository with a clear explanation of the research process. This could be in the audio-visual archive or elsewhere, and in any case in all the sources which she has used a reference should appear to her project. A next researcher with similar interests can find references to her work and choose to build on her conclusions.

4 Summary and conclusion

This paper reflects on the factors impeding a fruitful collaboration between humanities scholars and ICT-developers which have been observed in the course of a number of past and current projects. Special attention is paid to scholarly use of audio-visual collections. The line of thought is that if the potential for e-research and methodological innovation of the humanities is to materialize, the different mind-sets of ICT-researchers, archivists and scholars in the humanities should be geared to one another. In our view the road towards e-research as a welcome intervention in current practices, rather than as an imposed technology-driven paradigm, is through investments in education and training. Humanities scholars, and ICT-developers and students should all learn about the principles, challenges and biases of each other's discipline, especially in the preparatory phases of joint research projects. Collaboration according to principles of co-development is the second pillar. This entails early involvement of commercial software-developers, early investigations on tools that are already in use by scholars, and improvement of digital practices that have already been incorporated in existing practices.

Acknowledgements

We are grateful for financial support from the EU FP7 project AXES ICT-269980, the projects Infiniti and SEALINCMedia in the Dutch national FES program COMMIT, and the oral history project Verteld Verleden funded under the scheme Digitalisering met Beleid.

References

- Beaulieu, A. & Wouters, P. (2009). E-research as Intervention. In: *e-Research: Transformations in Scholarly Practice*, edited by N. Jankowski, (pp.54-69). New York: Routledge.
- Ceolin, D., van Hage, W., and Fokkink, W. (2010). A trust model to estimate the quality of annotations using the web. In: *Proceedings Web Science Conference 2010 - WebSci'10*. Raleigh, North Carolina.
- Goldman, J., Renals, S., Bird, S., de Jong, F. M. G., Federico, M., Fleischhauer, C., Kornbluh, M., Lamel, L., Oard, D. W., Stewart, C., & Wright, R. (2005). Accessing the spoken word. *International Journal on Digital Libraries*, 5(4), 287-298.
- Heuvel, W. van den (2010). Expert search for radio and television: a case study amongst Dutch broadcast professionals. In: *Proceedings of the 8th International Interactive Conference on Interactive TV&Video*. Tampere (pp. 47-50).
- Huurnink, B., Hollink, L., van den Heuvel, W., and de Rijke, M. (2010). Search behavior of media professionals at an audio-visual archive: A transaction log analysis. In: *Journal of the American Society for Information Science and Technology*, 61(6), 1180-1197.
- Jong, F.M.G. de, Oard, D.W., Heeren, W.F.L., Ordelman, R.J.F. (2008). Access to recorded interviews: A research agenda, *Journal on Computing and Cultural Heritage (JOCCH)*, 1(1), 3-29. ACM, New York.
- Ordelman, R.J.F., Heeren, W.F.L., Huijbregts, M.A.H., Jong, F.M.G. de, Hiemstra, D. (2009). Towards Affordable Disclosure of Spoken Heritage Archives, In: *Journal of Digital Information*, M.A. Larson, K. Fernie and J. Oomen (eds), 10(6), 17-33.
- RIN (2011). Reinventing research? Information practices in the humanities. A Research Information Network Report. Cf. <http://www.rin.ac.uk/>
- Scagliola, S. (2010). Speech Retrieval for Interview Data: Technology Push versus Academic Demand. In: *Proceedings of the 4th international workshop on Searching Spontaneous Conversational Speech*, at ACM Multimedia 2010, Florence (pp. 1-2). ACM, New York.